STAT 5443, HW3

Ariel Mundo

2020-10-24

PROBLEM 1

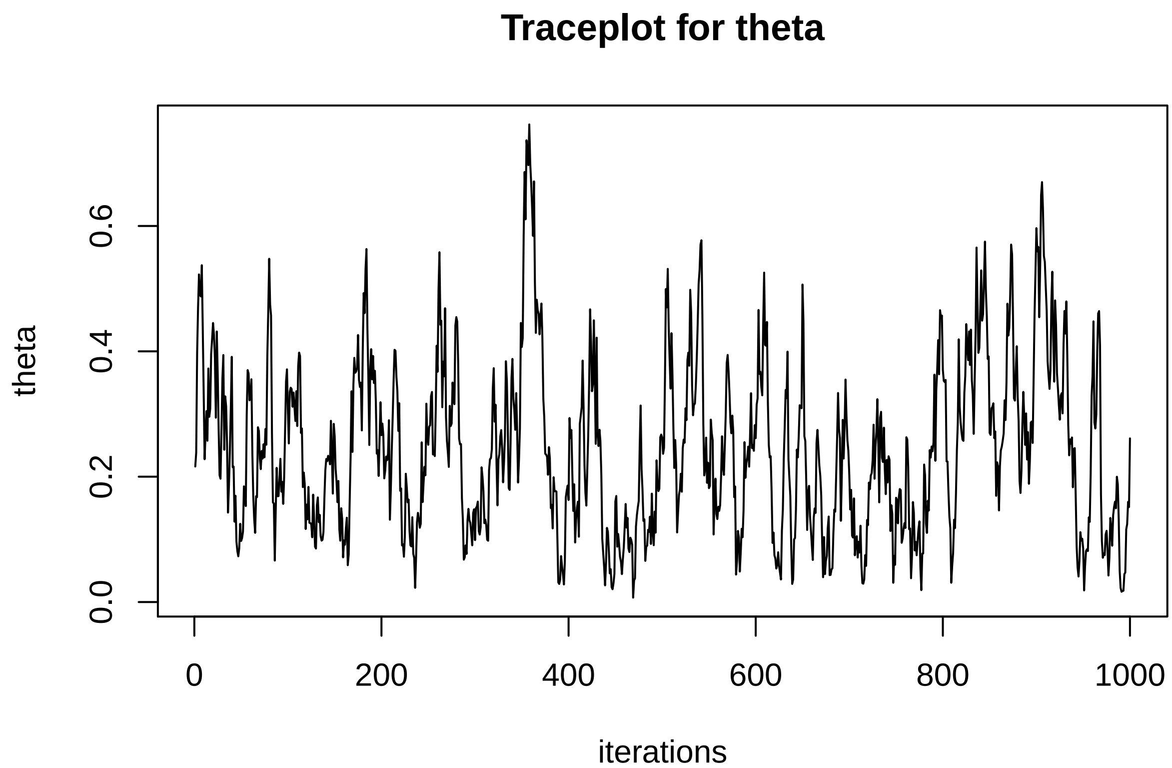
GIBBS SAMPLER

This section contains parts 1-4 from PROBLEM 1.

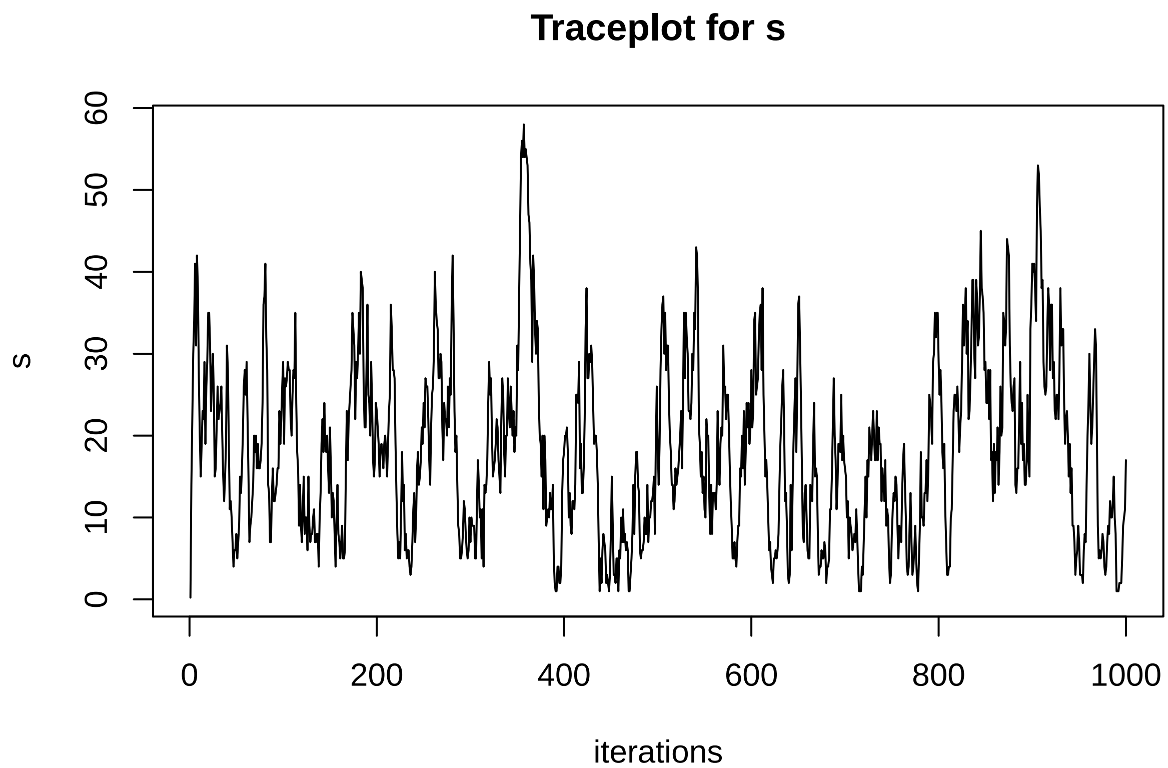
```
library(plyr)
```

```
## Warning: package 'plyr' was built under R version 3.5.3
```

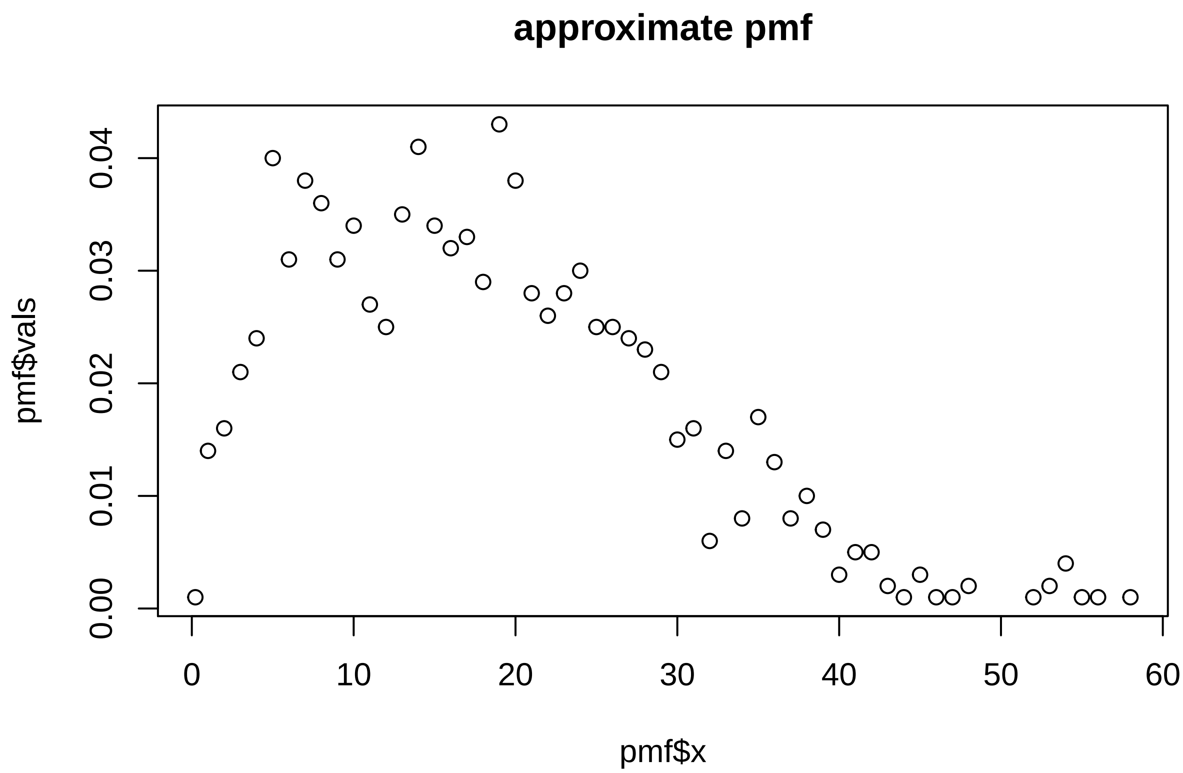
```
rm(list = ls())
set.seed(123)
alpha=2
beta_o=6.4
n=74
s_1=16
niter=1000
theta_init=s_1/n
s_init=s_1/n
  theta=rep(NA,niter)
  s=rep(NA,niter)
  theta[1] = theta_init
  s[1]=s_init
  for(i in 2:niter){
    s[i] = rbinom(1,size=n,prob=theta[i-1])
    theta[i] = rbeta(1,alpha+s[i],beta_o+n-s[i])
  }
  pmf<-count(s)
  pmf$vals<-pmf$freq/niter

plot(theta,type="l", main='Traceplot for theta',xlab='iterations')
```



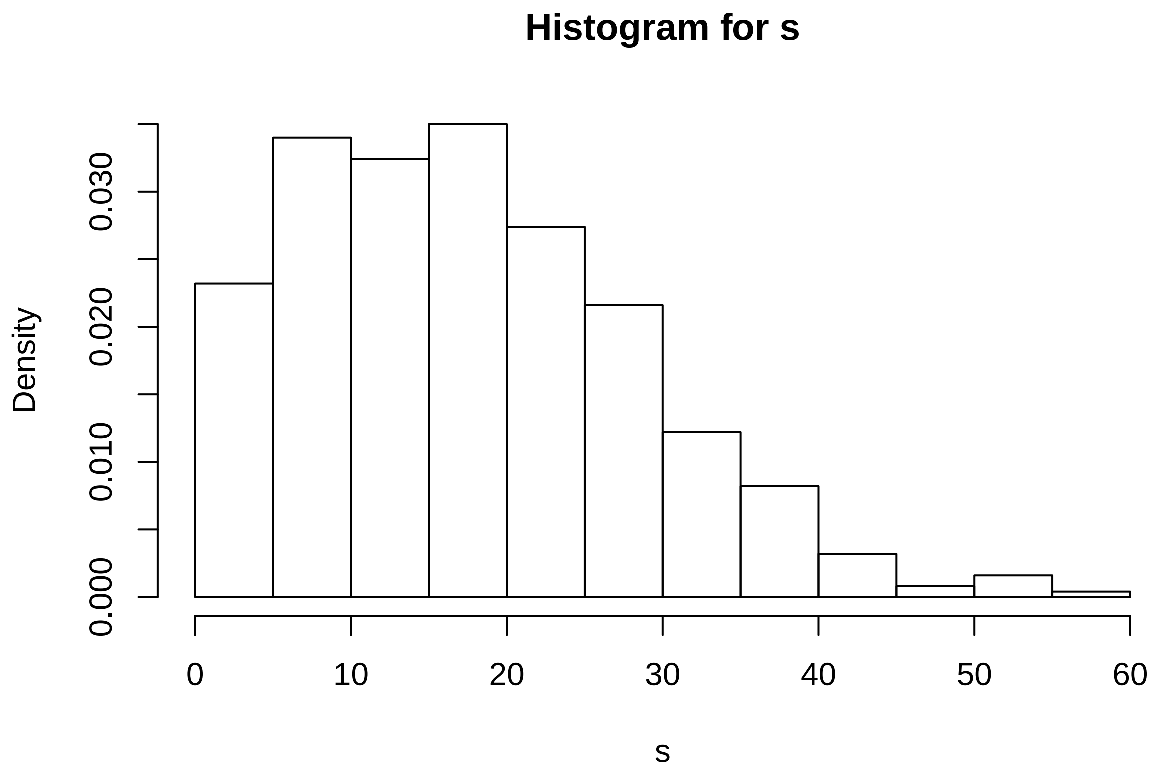

```
plot(s,type="l", main='Traceplot for s',xlab='iterations')
```




```
plot(x=pmf$x,y=pmf$vals, main='approximate pmf')
```




```
hist(s,freq=FALSE, main='Histogram for s')
```




```r
print(paste('posterior median of theta:',median(theta), 'MLE estimate',s_1/n))
```

```
## [1] "posterior median of theta: 0.230158313795905 MLE estimate 0.216216216216216"
```

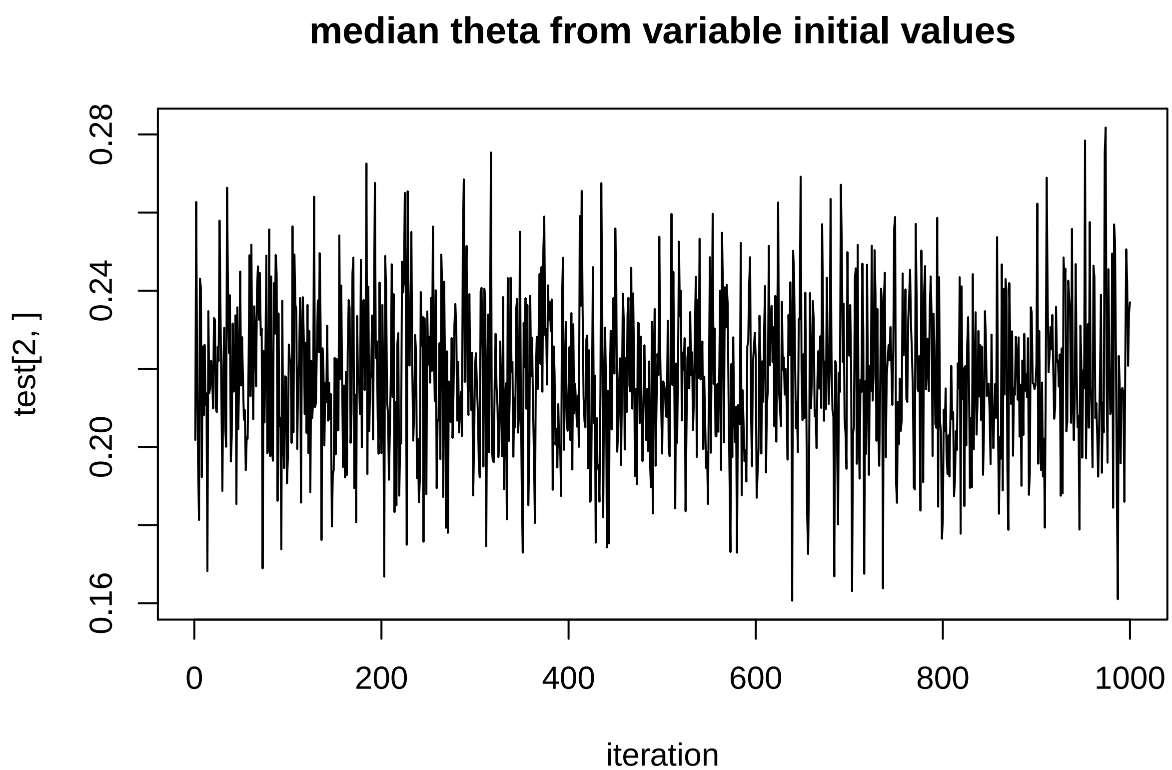
The estimated posterior mean of θ is indeed to the maximum likelihood estimate of s/n .

PART 5 OF PROBLEM 1

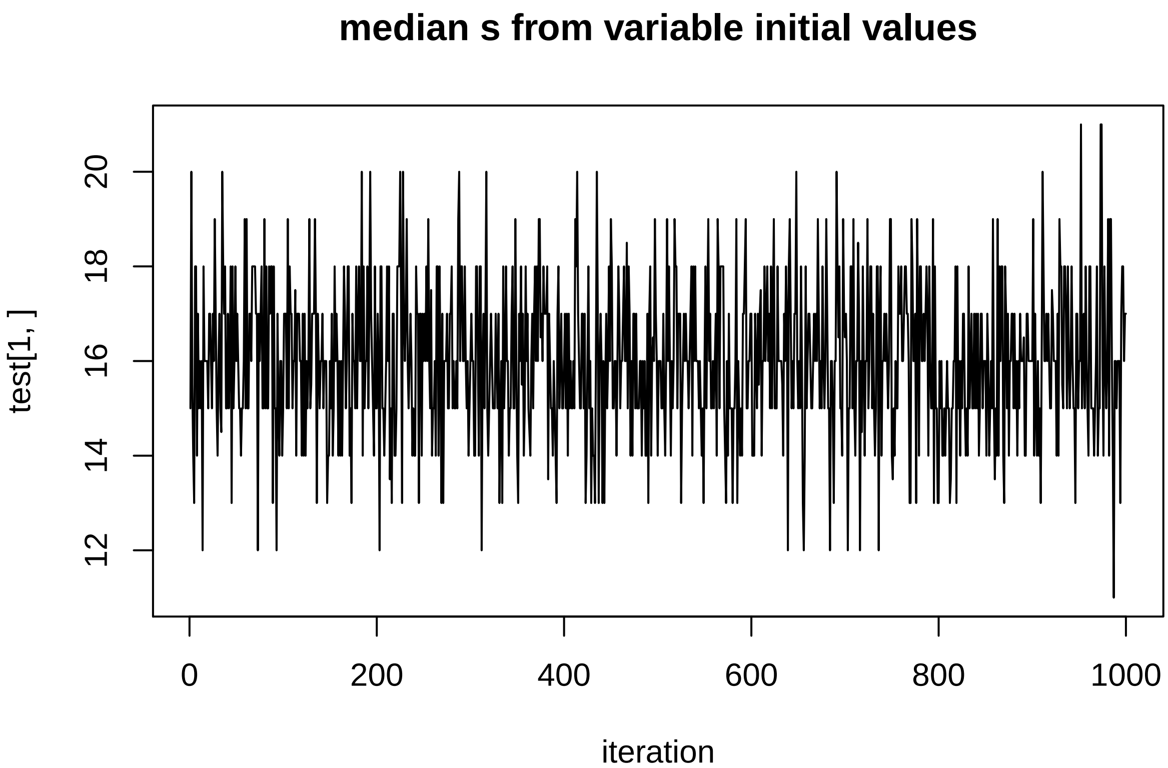
This section measures the sensitivity of the posterior median to the choice of initial values. For this part, the previous Gibbs sampler will be transformed in a function, and the initial ratio s/n will be randomized and the median s and θ will be plotted for all iterations.

```r
binomial_beta_gibbs<-function(theta_init,s_init,iterations){
alpha=2
beta_o=6.4
n=74
  theta_init=s_1/n
s_init=s_1/n
  theta=rep(NA,iterations)
  s=rep(NA,iterations)
  theta[1] = theta_init
  s[1]=s_init
  for(i in 2:iterations){
    s[i] = rbinom(1,size=n,prob=theta[i-1]) #conditional for s
    theta[i] = rbeta(1,alpha+s[i],beta_o+n-s[i]) #conditional for theta
  }
  pmf<-count(s)
  pmf$vals<-pmf$freq/niter
```

```
 return(cbind(median(s),median(theta)))
}
#v<-seq(0,10,0.1)
v<-runif(1000)
test<-matrix(NA,2,length(v))
for (j in 1:length(v)){
test[,j]<-binomial_beta_gibbs(v[j],v[j],1000)
}
plot(test[2,], type='l', xlab='iteration',main='median theta from variable initial values') #plotting t
```




```
plot(test[1,], type='l', xlab='iteration',main='median s from variable initial values') #plotting the m
```



PROBLEM 2

The pmf in this case is the result of the combination of each individual distribution:

$\binom{n}{s}\theta^s(1-\theta)^{n-s}\frac{\Gamma(\alpha_0+\beta_0)}{\Gamma(\alpha_0)\Gamma(\beta_0)}\theta^{\alpha_0-1}(1-\theta)^{\beta_0-1}\left[\frac{\exp(-\lambda)\lambda^n}{n!})\right]$

This can be re-written as: $\frac{n!}{s!(n-s!)}\theta^s(1-\theta)^{n-s}\frac{\Gamma(\alpha_0+\beta_0)}{\Gamma(\alpha_0)\Gamma(\beta_0)}\theta^{\alpha_0-1}(1-\theta)^{\beta_0-1}\left[\frac{\exp(-\lambda)\lambda^n}{n!})\right]$

For $\theta|s,n \propto \theta^{\alpha_0+s-1}(1-\theta)^{n-s+\beta_0-1} \propto Beta(\alpha_0+s, n-s+\beta_0)$

For $s|\theta,n \propto \binom{n}{s}\theta^s(1-\theta)^{n-s} \propto Binomial(n,\theta)$

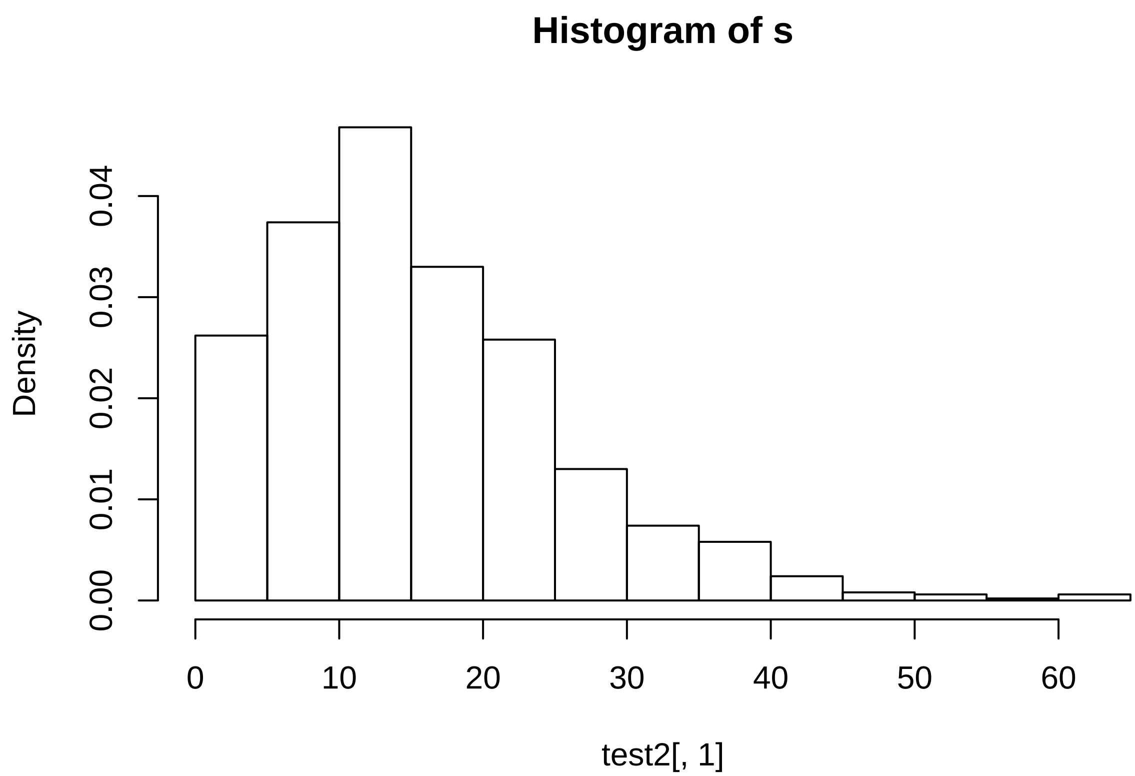
Therefore the two terms $n!$ cancel out and the exponential can be taken out as it is a constant.

Writing the conditional distribution of n by considering only the terms that include it and omitting constant terms: $f(n|\theta,s) \propto \frac{\lambda^n(1-\theta)^{n-s}}{(n-s)!}$

Which, considering a constant $\exp(-\lambda(1-\theta))$ means that $(n-s) \propto Poisson(\lambda(1-\theta))$

```
n_1=74
binomial_beta_gibbs_poisson<-function(n_init,theta_init,s_init,iterations){
alpha=2
beta_o=6.4
lambda=64
  theta_bgp=rep(NA,iterations)
  s_bgp=rep(NA,iterations)
  n_bgp=rep(NA,iterations)
```

```
    theta_bgp[1] = theta_init
    s_bgp[1]=s_init
    n_bgp[1]=n_init
  for(i in 2:iterations){
    s_bgp[i] = rbinom(1,size=n_bgp[i-1],prob=theta_bgp[i-1])
    theta_bgp[i] = rbeta(1,alpha+s_bgp[i],beta_o+n_bgp[i-1]-s_bgp[i])
    n_bgp[i]=s_bgp[i]+rpois(1,lambda*(1-theta_bgp[i])) #conditional for n
  }
  #pmf<-count(s)
  #pmf$vals<-pmf$freq/niter
 return(cbind(s_bgp,theta_bgp,n_bgp)) #storing the values
}

test2<-binomial_beta_gibbs_poisson(10,s_1/n_1,s_1/n_1,1000)
hist(test2[,1],freq=FALSE, main='Histogram of s')
```

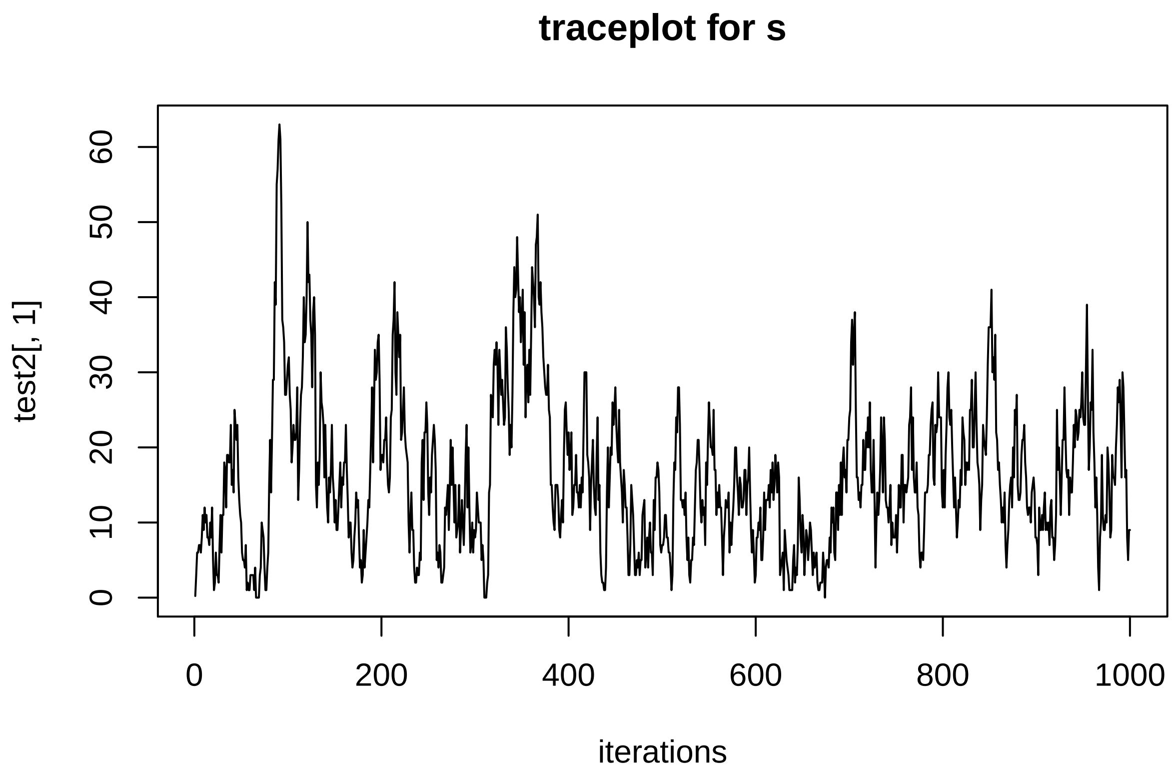



```
print(paste('Median theta of Binomial-Beta-Poisson',median(test2[,2]),'The posterior median is similar
```

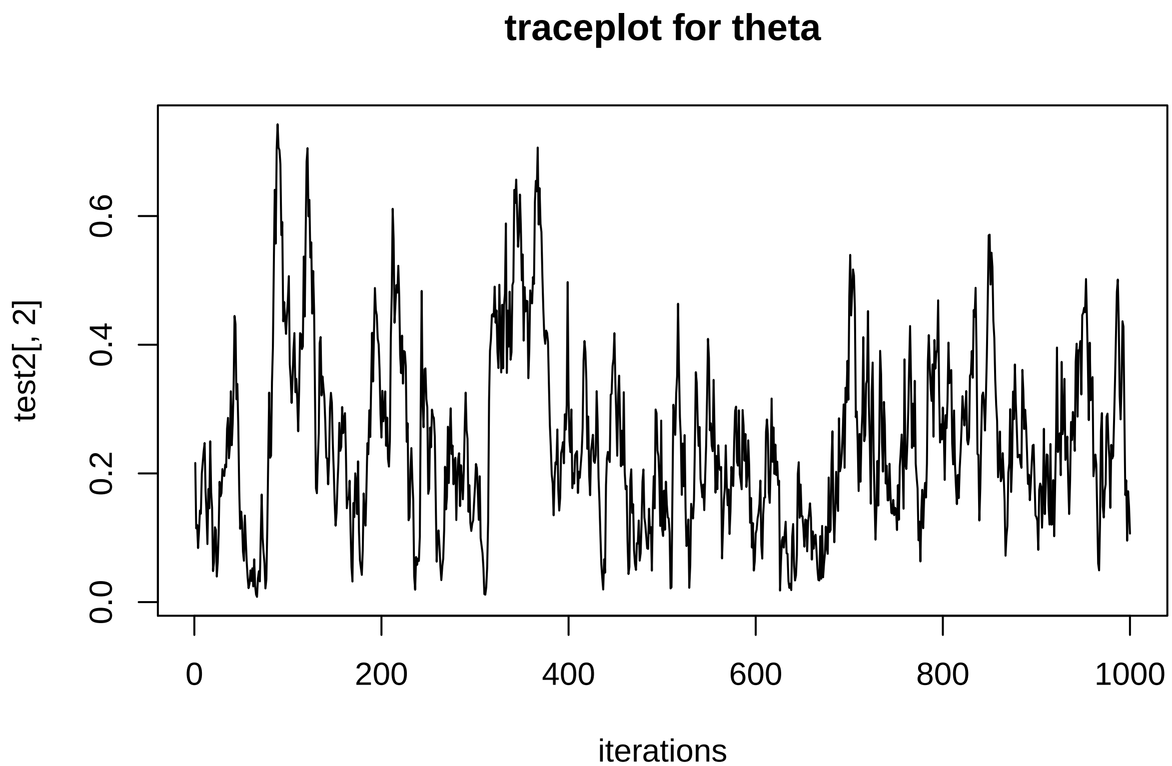
```
## [1] "Median theta of Binomial-Beta-Poisson 0.227290381804581 The posterior median is similar to the
```

Regarding convergence, plotting the traceplot for s and θ:

```
plot(test2[,1], type='l',main='traceplot for s', xlab='iterations')
```




```
plot(test2[,2], type='l',main='traceplot for theta', xlab='iterations')
```



traceplot for theta
0.0
0.2
0.4
0.6
test2[, 2]
0
200
400
600
800
1000
iterations